Let $X_1, \ldots, X_n$ be a random sample from a normal distribution $N(\mu_X, \sigma^2)$, and $Y_1, \ldots, Y_m$ be a random sample from a normal distribution $N(\mu_Y, \sigma^2)$, with $\mu_X, \mu_Y$ and $\sigma$ unknown. We assume the $X$’s and $Y$’s are independent. Let $\overline{X}$ and $s_X$ denote the sample mean and sample standard deviation of the $X_1, \ldots, X_n$, respectively. (Similar meaning for $\overline{Y}$ and $s_Y$.) Define the pooled sample variance as

$$s_p = \frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}.$$  

We want to test the hypotheses $H_0 : \mu_X - \mu_Y = \Delta$ v.s. $H_A : \mu_X - \mu_Y > \Delta$. We use the test statistic

$$T = \frac{(\overline{X} - \overline{Y}) - \Delta}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}}$$

and rejection region $T > c$.

a. Show that the significance level of this test is $1 - F_{n+m-2}(c)$, where $F_{n+m-2}$ is the cdf of the $t$-distribution with degrees of freedom $n + m - 2$, that is, $F_{n+m-2}(t) = P(U \leq t)$ with $U \sim t_{n+m-2}$.

b. Show that if the test is required to have significance level $\alpha$, then $c = t_{1-\alpha, n+m-2}$. Here $t_{1-\alpha, n+m-2}$ satisfies $P(U \leq t_{1-\alpha, n+m-2}) = 1 - \alpha$.

c. Show that if the sample gives $T = t$, then the $p$-value in this case is $1 - F_{n+m-2}(t)$.

d. Assume $\sigma$ is known. Show that if in fact $\mu_X - \mu_Y = \delta$ with $\delta > \Delta$, then the power of the test is $1 - \Phi \left( c + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right)$. If the significance level is $\alpha$, then the power is $1 - \Phi \left( z_{1-\alpha} + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right)$.

a. Since the significance level is the probability of making a type I error, we have

$$\alpha = P(\text{reject } H_0 \mid H_0) = P(T > c \mid H_0) = P \left( \frac{(\overline{X} - \overline{Y}) - \Delta}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} > c \mid H_0 \right) = 1 - F_{n+m-2}(c),$$
since
\[
\frac{(\bar{X} - \bar{Y}) - \Delta}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2} \text{ under } H_0.
\]

b. From the above derivation, we have
\[
\alpha = 1 - F_{n+m-2}(c), \quad \text{or} \quad F_{n+m-2}(c) = 1 - \alpha.
\]
By the definition of \( t_{1-\alpha,n+m-2} \), we have
\[
c = t_{1-\alpha,n+m-2}.
\]

c. Since the \( p \)-value is the probability of making type I error when we reject \( H_0 \) based on the sample, we have (by following a similar argument as in (a))
\[
p = P(\text{reject } H_0 \text{ based on the sample } |H_0) = P(T > t|H_0)
\]
\[
= P\left( \frac{(\bar{X} - \bar{Y}) - \Delta}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} > t \bigg| H_0 \right)
\]
\[
= 1 - F_{n+m-2}(t),
\]

since
\[
\frac{(\bar{X} - \bar{Y}) - \Delta}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2} \text{ under } H_0.
\]

d. Note that here we assume \( \sigma \) is known. Correspondingly, we will change our test statistic to
\[
T = \frac{(\bar{X} - \bar{Y}) - \Delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}},
\]
Therefore, we have
\[
\frac{(\bar{X} - \bar{Y}) - \Delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0,1) \text{ under } H_0,
\]
instead of
\[
\frac{(\bar{X} - \bar{Y}) - \Delta}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2} \text{ under } H_0.
\]
A similar argument as in (a) shows that

\[
\alpha = P(\text{reject } H_0 | H_0) = P(Z > c | H_0) = P \left( \frac{(\bar{X} - \bar{Y}) - \Delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} > c | H_0 \right) = 1 - \Phi(c),
\]

since

\[
\frac{(\bar{X} - \bar{Y}) - \Delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0, 1) \text{ under } H_0,
\]

Thus \( c = z_{1-\alpha} \). The probability of making type II error is

\[
\beta = P(\text{accept } H_0 | H_A) = P(Z \leq c | H_A)
\]

\[
= P \left( \frac{(\bar{X} - \bar{Y}) - \Delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \leq c \bigg| H_A \right)
\]

\[
= P \left( \frac{(\bar{X} - \bar{Y}) - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \leq c + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \bigg| H_A \right)
\]

\[
= P \left( Z \leq c + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right) \quad Z \sim N(0, 1)
\]

\[
= \Phi \left( c + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right).
\]

Furthermore, the power of the test is

\[
\text{power} = 1 - \beta = 1 - \Phi \left( c + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right).
\]

When the significance level is \( \alpha \), we derived from part (b) that

\[
c = z_{1-\alpha}.
\]

Thus, by combining the above two equations, the power of this test
2. Low-back pain (LBP) is a serious health problem in many industrial settings. The article “Isodynamic Evaluation of Trunk Muscles and Low-Back Pain Among Workers in a Steel Factory” (Ergonomics, 1995: 2107-2117) reported the accompanying summary data on lateral range of motion (degrees) for a sample of workers without a history of LBP and another sample with a history of this malady.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Sample Size</th>
<th>Sample Mean</th>
<th>Sample SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>No LBP</td>
<td>28</td>
<td>91.5</td>
<td>5.5</td>
</tr>
<tr>
<td>LBP</td>
<td>31</td>
<td>88.3</td>
<td>7.8</td>
</tr>
</tbody>
</table>

We assume that the lateral range of motion under both conditions are normally distributed and have a common (unknown) standard deviation $\sigma$.

a. Calculate a 90% confidence interval for the difference between population mean extent of lateral motion for the two conditions.

b. Does the data suggest that population mean lateral motion differs for the two conditions?

Let $X$ denote the sample for “no LBP” and $Y$ for “LBP”. Thus the sample size, sample mean, and sample standard deviation for $X$ and $Y$ are

- $\bar{x} = 91.5$, $s_X = 5.5$, $n = 28$,
- $\bar{y} = 88.3$, $s_Y = 7.8$, $m = 31$,

respectively. Let $\mu_X$ and $\mu_Y$ denote the population mean extent of lateral motion for the two conditions, respectively.
(a) A 100(1 - \alpha)\% confidence interval for \(\mu_X - \mu_Y\) is given by

\[
\left( (\bar{x} - \bar{y}) - t_{1-\frac{\alpha}{2}, n+m-2} \cdot s, \ (\bar{x} - \bar{y}) - t_{1-\frac{\alpha}{2}, n+m-2} \cdot s \right),
\]

where \(s\) is the estimated standard error (deviation) of \(X - Y\):

\[
s = s_p \sqrt{\frac{1}{n} + \frac{1}{m}},
\]

and \(s_p\) is the pooled sample standard deviation:

\[
s_p = \sqrt{\frac{(n-1)s_X^2 + (m-1)s_Y^2}{n + m - 2}}.
\]

Since \(\alpha = 0.1(= 1 - 0.9)\), we have \(t_{1-\frac{0.1}{2}, 28 + 31 - 2} = t_{0.95,57} = 1.672\) (if we use TABLE 4 of the textbook, we can use \(t_{0.95,60} = 1.671\) to approximate). Moreover,

\[
s_p = \sqrt{\frac{(28 - 1) \cdot 5.5^2 + (31 - 1) \cdot 7.8^2}{28 + 31 - 2}} = 6.808,
\]

and

\[
s = 6.808 \cdot \sqrt{\frac{1}{28} + \frac{1}{31}} = 1.775.
\]

Therefore, a 90% confidence interval for \(\mu_X - \mu_Y\) is

\[
(91.5 - 88.3) - 1.672 \cdot 1.775, \ (91.5 - 88.3) + 1.672 \cdot 1.775
\]

\[= (0.232, 6.168).\]

If we use \(t_{0.95,57} \approx t_{0.95,60} = 1.671\), then the C.I. is \((0.234, 6.166)\).

b. The hypotheses are

\[H_0 : \mu_X = \mu_Y \quad H_A : \mu_X \neq \mu_Y.\]

The test statistic is

\[
T = \frac{\bar{X} - \bar{Y}}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}},
\]

and the rejection region is of the form \(\{|T| > c\}\). Since for this sample

\[
T = \frac{91.5 - 88.3}{1.775} = 1.803
\]

and \(T \sim t_{57}\), the p-value is

\[
p = 2[1 - F_{57}(1.803)] = 0.077.
\]
Thus if our significance level is $\alpha = 5\%$, we will not reject $H_0$, that is, we will conclude the population mean lateral motion does not differ for the two conditions. However, if we use $\alpha = 10\%$, then we will reject $H_0$, that is, we will conclude that population mean lateral motion differs for the two conditions.

If we use TABLE 4, we will use $df = 60$ to approximated $df = 57$. In this case, we have $t_{0.95,60} = 1.671 < 1.803 < 2.000 = t_{0.975,60}$. Thus the $p$-value satisfies

$$2 \cdot (1 - 0.975) = 0.05 < p < 0.1 \Rightarrow 2 \cdot (1 - 0.95),$$

which means we reject $H_0$ at $\alpha = 0.1$ but not at $\alpha = 0.05$.

Or equivalently, we can use rejection region $\{T > 1.671 = t_{0.95,60}\}$ corresponding to $\alpha = 0.1$, and $\{T > 2.000 = t_{0.975,60}\}$ corresponding to $\alpha = 0.05$ in order to reach the same conclusion.

3. The article “The Effects of a Low-Fat, Plant-Based Dietary Intervention on Body Weight, Metabolism, and Insulin Sensitivity in Postmenopausal Women” (Amer. J. of Med., 2005: 991-997) reported on the results of an experiment in which half of the individuals in a group of 64 postmenopausal overweight women were randomly assigned to a particular vegan diet, and the other half received a diet based on National Cholesterol Education Program guidelines. The sample mean decrease in body weight for those on the vegan diet was 5.8 kg, and the sample SD was 3.2, whereas for those on the control diet, the sample mean weight loss and standard deviation were 3.8 and 2.8, respectively. Does it appear the true average weight loss for the vegan diet exceeds that for the control diet by more than 1 kg? Assume the weight loss under both diet are normally distributed with an unknown variance $\sigma^2$. Carry out an appropriate test of hypotheses at significance level .05 based on calculating a $p$-value.
Let \( X \) denote the sample for the group with the vegan diet and \( Y \) for the group with control diet. Thus the sample size, sample mean, and sample standard deviation for \( X \) and \( Y \) are

\[
\bar{x} = 5.8, \quad s_X = 3.2, \quad n = 32,
\]
\[
\bar{y} = 3.8, \quad s_Y = 2.8, \quad m = 32,
\]
respectively. Let \( \mu_X \) and \( \mu_Y \) denote the population average weight loss for the two diets, respectively. They hypotheses are

\[
H_0 : \mu_X - \mu_Y = 1 \quad H_A : \mu_X - \mu_Y > 1.
\]

The test statistic is

\[
T = \frac{(\bar{X} - \bar{Y}) - 1}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}},
\]
and the rejection region is of the form \( \{T > c\} \). For this sample, we have

\[
s_p = \sqrt{(32 - 1) \cdot 3.2^2 + (32 - 1) \cdot 2.8^2} = 3.007,
\]

and

\[
T = \frac{(5.8 - 3.8) - 1}{3.007 \cdot \sqrt{\frac{1}{32} + \frac{1}{32}}} = 1.330.
\]

Since \( T \sim t_{62} \), the \( p \)-value is

\[
p = 1 - F_{62}(1.330) = 0.094.
\]

Therefore, we will not reject \( H_0 \) at significance level \( \alpha = 0.05 \), that is, there is no significant evidence that the true average weight loss for the vegan diet exceeds that for the control diet by more than 1 kg.

We can also find the rejection region according to \( \alpha = 0.05 \): \( \{T > t_{0.95,62} = 1.670\} \). The same conclusion will be made since 1.330 < 1.670.

If we use TABLE 4 of the textbook, we will use \( df = 60 \) to approximate \( df = 62 \). In this case, we will have \( t_{0.0.90,60} = 1.296 < 1.330 < 1.671 = t_{0.95,60} \) which implies that \( \alpha = 0.05 < p < 0.10 \), and \( t_{0.95,60} = 1.671 > 1.330 \).

4. [§11-15] Suppose that \( n \) measurements are to be taken under a treatment condition and another \( n \) measurements are to be taken independently under a control condition. It is thought that the standard deviation of a single observation is about 10 under both conditions. How large should \( n \) be so that a 95% confidence interval for \( \mu_X - \mu_Y \) has a width of 2? Use the normal distribution rather than the \( t \) distribution, since we assume \( \sigma = 10 \).
When $\sigma$ is known, we use the fact

$$X - Y \sim N \left( \mu_X - \mu_Y, \sigma \sqrt{\frac{1}{n} + \frac{1}{m}} \right).$$

Correspondingly, a $100(1 - \alpha)\%$ confidence interval for $\mu_X - \mu_Y$ is given by

$$\left( \bar{X} - \bar{Y} - z_{1-\alpha/2} \cdot \sigma \sqrt{\frac{1}{n} + \frac{1}{m}}, \bar{X} - \bar{Y} + z_{1-\alpha/2} \cdot \sigma \sqrt{\frac{1}{n} + \frac{1}{m}} \right).$$

Thus the width of the confidence interval is

$$w = 2z_{1-\alpha/2} \cdot \sigma \sqrt{\frac{1}{n} + \frac{1}{m}}.$$

In this case, we have $m = n$. Therefore

$$w = 2z_{1-\alpha/2} \cdot \sigma \sqrt{\frac{2}{n}}.$$

In order to reach a desired width of confidence intervals with given confidence level, we need sample size to satisfy

$$n = \left( \frac{2\sqrt{2} \cdot z_{1-\alpha/2} \cdot \sigma}{w} \right)^2.$$

Since $\alpha = 0.05(= 1 - 0.95)$, we have $z_{0.975} = 1.96$. In this case, $\sigma = 10$ and $w = 2$. Thus,

$$n = \left( \frac{2\sqrt{2} \cdot 1.96 \cdot 10}{2} \right)^2 = 768.32 \approx 769.$$

Note that sample size $n$ must be a positive integer. And we typically round up when the calculated value is not an integer. The reason is that we want to guarantee that the width is no more than the given value and the confidence level is no less than the given value.

Appendix

1. Assume the same setting as in Problem 1. In this case we want to test the hypotheses $H_0: \mu_X - \mu_Y = \Delta$ v.s. $H_A: \mu_X - \mu_Y < \Delta$. We use the test statistic

$$T = \frac{(X - Y) - \Delta}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}}.$$

and rejection region $T < c$. 

8
a. The significance level of this test is $F_{n+m-2}(c)$.

b. If the test is required to have significance level $\alpha$, then $c = t_{\alpha,n+m-2}$.

c. If the sample gives $T = t$, then the $p$-value in this case is $F_{n+m-2}(t)$.

d. Assume $\sigma$ is known. If in fact $\mu_X - \mu_Y = \delta$ with $\delta < \Delta$, then the power of the test is $\Phi \left( c + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right)$. If the significance level is $\alpha$, then the power is $\Phi \left( z_\alpha + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right)$.

1. Assume the same setting as in Problem 1. In this case we want to test the hypotheses $H_0: \mu_X - \mu_Y = \Delta$ v.s. $H_A: \mu_X - \mu_Y \neq \Delta$. We use the test statistic

$$ T = \frac{(\bar{X} - \bar{Y}) - \Delta}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}}, $$

and rejection region $|T| > c$.

a. The significance level of this test is $2(1 - F_{n+m-2}(c))$.

b. If the test is required to have significance level $\alpha$, then $c = t_{\frac{\alpha}{2},n+m-2}$.

c. If the sample gives $T = t$, then the $p$-value in this case is $2(1 - F_{n+m-2}(|t|))$.

d. Assume $\sigma$ is known. If in fact $\mu_X - \mu_Y = \delta$ with $\delta \neq \Delta$, then the power of the test is $1 - \Phi \left( c + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right) + \Phi \left( -c + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right)$. If the significance level is $\alpha$, then the power is

$$ 1 - \Phi \left( z_{1-\frac{\alpha}{2}} + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right) + \Phi \left( -z_{1-\frac{\alpha}{2}} + \frac{\Delta - \delta}{\sigma \sqrt{\frac{1}{n} + \frac{1}{m}}} \right). $$